




Scholars' Perceptions of Relevance in Bibliography-Based People Recommender System

Ekaterina Olshannikova^{1,*} , Thomas Olsson¹, Jukka Huhtamäki¹ & Peng Yao¹
¹*Tampere University, Tampere, Finland (E-mail: ekaterina.olshannikova@tuni.fi;
E-mail: thomas.olsson@tuni.fi; E-mail: jukka.huhtamaki@tuni.fi; E-mail: peng.yao@tuni.fi)*

Abstract. Collaboration and social networking are increasingly important for academics, yet identifying relevant collaborators requires remarkable effort. While there are various networking services optimized for seeking similarities between the users, the scholarly motive of producing new knowledge calls for assistance in identifying people with complementary qualities. However, there is little empirical understanding of how academics perceive relevance, complementarity, and diversity of individuals in their profession and how these concepts can be optimally embedded in social matching systems. This paper aims to support the development of diversity-enhancing people recommender systems by exploring senior researchers' perceptions of recommended other scholars at different levels on a similar–different continuum. To conduct the study, we built a recommender system based on topic modeling of scholars' publications in the DBLP computer science bibliography. A study of 18 senior researchers comprised a controlled experiment and semi-structured interviewing, focusing on their subjective perceptions regarding relevance, similarity, and familiarity of the given recommendations, as well as participants' readiness to interact with the recommended people. The study implies that the homophily bias (behavioral tendency to select similar others) is strong despite the recognized need for complementarity. While the experiment indicated consistent and significant differences between the perceived relevance of most similar vs. other levels, the interview results imply that the evaluation of the relevance of people recommendations is complex and multifaceted. Despite the inherent bias in selection, the participants could identify highly interesting collaboration opportunities on all levels of similarity.

Key words: People recommender systems, Social matching applications, Expert search, Homophily, Diversity, Research collaboration, Social networking, Bibliography data analysis

1. Introduction

In scholarly work, collaboration has become a normative form of knowledge production. Researchers across the social sciences broadly concur that collaboration is the best path to solving complex problems and achieving exceptional results (Frydinger et al. 2013). Collaboration is promoted as a means of cultivating quality, enhanced resource utilization, and high impact (Hsiehchen et al. 2015). In

science and patenting, a substantial shift toward collective work has been found across scientific disciplines and business domains (Wuchty et al. 2007; Börner et al. 2010). In academic research, collaboration takes place on a dyadic level between individuals, amongst research teams, as well as within international consortia. However, identifying new suitable candidates for academic collaboration requires high investment in social networking, and the disciplinary structures can prevent unexpected combinations of individuals.

Following this trend of the increasing importance of collaboration, supporting social networking and encouraging new social encounters have become central design goals in the HCI & CSCW communities. Prior research on so-called social matching (Terveen and McDonald 2005) has particularly looked into people recommender systems (Tsai and Brusilovsky 2016; Guy and Pizzato 2016) and opportunistic matching applications (Mayer et al. 2015a; Mayer et al. 2016) that aim to enable identification of new relevant connections, some of them employing playful approaches and gamification (Paasovaara et al. 2016). There are also prototypes of people recommender systems that specifically aim to match scholars: for instance, expert finding systems (Vassileva et al. 2003; Beham et al. 2010), or event-based mobile applications like *‘Find & Connect’* developed by Chin et al. (2014) and experimented at the UbiComp 2011 conference. Considering the rich publication data available in online repositories, prior research has looked into bibliography analysis methods for recommender systems, e.g., DBLP¹-based systems for researchers (Zaiane et al. 2007).

However, the majority of professional matching systems tend to utilize a similarity-maximizing approach, providing recommendations of like-minded others with similar interests. In this regard, Yuan and Gay (2006) deliberate that homogeneity produces both positive and negative effects on interpersonal communication, community formation, and knowledge work – *“homophily not only unifies, it also divides a network”*. On the one hand, collaboration within a group of people with shared interests can contribute to a safe and trustworthy work environment, enabling cohesive team spirit and ease of communication. On the other hand, it has been found that researching and cooperating with diverse individuals is essential in tasks that aim to create new knowledge (Mollica et al. 2003). Prior work emphasizes that an insightful dialogue between diverse actors can build social capital (Burt 2017) by increasing awareness about external knowledge groups and bridging polarized intellectual communities towards abounding knowledge sharing and idea creation (Argote and Ophir 2002). While research on diversifying item recommendations (Adamopoulos and Tuzhilin 2015; Castells et al. 2015) is gaining interest, few attempts have been made to match people based on diversity (Rajagopal et al. 2017).

Additionally, the evaluation of people recommender systems and matching applications is geared toward the assessment of algorithm effectiveness, with little focus on user perceptions. Although there is well-established research on user-centered evaluation of content recommender systems (Knijnenburg et al. 2012;

Pu et al. 2011), the choice of potential collaborators is significantly different and, therefore, requires contextually operationalized evaluation metrics. Considering the diverse needs of scholars, it is essential to pay attention to subjective perceptions regarding the recommended people and carefully conceptualize factors such as perceived relevance and willingness to follow-up on the recommendations.

To enable the gathering of such data, we developed a simple DBLP-based people recommender system that provides the user with recommendations of other scholars from three different levels of similarity regarding their publication history – low, moderate and high. With a user study that combines a controlled experiment and semi-structured interviewing, we address the following research questions: **(RQ1)** *What level of measured similarity of publication history is preferred in recommendations of potential collaborators?* **(RQ2)** *What specific needs and expectations scholars have in regard to seeking professional collaboration?*

The findings reveal an intriguing mismatch between scholars' intuitive behavior and deliberate intentions regarding potential academic collaboration. While the quantitative results demonstrate participants' general preference to most similar recommended people, the interview data brings up a variety of scholars' needs for connecting with cross-disciplinary and diverse people. Thus, the nature of the collaboration task might influence the perceived relevance of potential candidates, for example, regarding the complementarity of professional roles, skills, and expertise.

The contribution of this work is two-fold: (i) providing empirical findings on subjective perceptions of people recommendation relevance in the context of potential academic partnering, and (ii) presenting the qualitative account of academics' needs in collaboration and factors that might affect their decision in choosing partners. Furthermore, as a methodological contribution, we operationalize measures for subjective opinions on people recommendations in the context of professional academic collaboration.

2. Related work

While various disciplines have studied scientific collaboration in different ways, a general consensus is that collaboration is imperative for effective knowledge production. Bozeman et al. (2013) note that research collaboration is often limited to the notion of co-authorship, and criticize the assumption that cooperation is undoubtedly resulting in a knowledge product (e.g., scientific paper). In this article, we approach scholarly partnering practices, which are beyond co-authorship, and do not require an explicit valuable outcome. In fact, Bozeman and Corley (2004) define key motives for scientific collaboration and propose that the selection of collaborators can be driven by: (i) work ethic attribution and schedule compliance; (ii) shared nationality; (iii) need to mentor junior researchers; (iv) administration request or high reputation; (v) preceding collaboration experience, its quality and personality chemistry; (vi) complementarity of skills. Considering this breadth, in this article we adopt the broad definition by Bozeman et al.

(2013): “*collaboration is a social process whereby human beings pool their human capital for the objective of producing knowledge.*”

To emphasize novelty of our contribution and the research gap we cover the following topics. First, we discuss research on supporting social interaction and collaboration in the general context of conferences and introduce works on bibliography-based recommender systems. Then, we deliberate on the concepts of similarity and diversity. Finally, we provide an overview of existing user-centered evaluation metrics in recommender systems.

2.1. Computational support for matching scholars

Supporting experts finding is one of the crucial CSCW design goals to facilitate collaborative knowledge creation and dissemination (Ackerman and McDonald 1996). Over the last two decades, systems for supporting the conference experience have expanded from increasing people’s awareness of necessary information at the venue (e.g., schedule and contents) to facilitating social encounters, for example arranging meetings (Nishibe et al. 1998) and general enhancement of attendee interaction with the environment and other people (Dey et al. 1999). One reasonably common approach relates to location and proximity-based services for finding relevant connections (Kawakita et al. 2004; Cox et al. 2003). For instance, ‘*Find & Connect*’, a social networking mobile application developed by Chin et al. (2014) and experimented during the UbiComp 2011 conference, aims to provide the users with social recommendations based on physical proximity and similarity of interests. The results of the work reveal that users preferred acting on familiar recommended people or friends-of-friends, as well as those who have similar research interests.

There are also generic, platform-like services designed to help the attendees meet new people with shared interests (Zenk et al. 2014). An example of such tool is ‘*Confer*’ (Zhang et al. 2016) which has been tested and deployed in several HCI conferences. ‘*Conference Navigator*’ is another example and has gradually acquired new features and functions (Farzan and Brusilovsky 2007; Wongchokprasitti et al. 2010; Parra et al. 2012). For instance, the most recent version (Brusilovsky et al. 2017) – ‘*Conference Navigator 3*’ – is a community-based recommender system that by utilizing content-based and tag-based analysis methods provide the user with personalized suggestions about people and contents of a conference. The user can explore their research community through interactive social network visualization and connect with similar experts either before or during an event.

Another recent research proposes so-called ‘*Adaptive Conference Companion*’ (Arens-Volland and Naudet 2016) – a mobile application that aims to deliver personalized guidance for attendees of academic events. In addition to utilizing conference data and explicit user input, authors enhance profiling and match-making mechanism by extracting bibliographic database (DBLP, GoogleScholar)

and social media channels (LinkedIn, ResearchGate and MyScienceWork). They applied a term frequency-inverse document frequency (TF-IDF)² algorithm (Beel et al. 2016) for recommending most similar people and sessions in the scope of users' interests. The authors speculate that the majority of participants in the experiment were well-prepared for the conference and already had their schedule which matched with recommendation predictions of relevant content. Unfortunately, the authors did not discuss the effectiveness of a system from the social networking perspective, i.e., how people reacted to and perceived the relevance of people recommendations.

Another vein of research relates to recommending academic collaborations by utilizing bibliography data and social networks analysis and, thus, suggesting candidates with similar research interests. For instance, Kong et al. (2016) utilized topic clustering model to retrieve academic domains, calculate authors' features, and analyze academic collaboration networks. Another group of researchers (Li et al. 2014) explored co-authorship networks to identify relevant collaborators with an already existing academic tie. They observed that scholars' relationships are more complicated in a real-world setting and suggested that future work should go beyond existing co-authorship networks and consider matching people without established connections. Most recently, Hoang et al. (2017) proposes a new approach to calculate similarity with deep learning and experiment on DBLP and WiKiCFP databases. Sie et al. (2012) designed a system for recommending future co-authors that utilizes co-authorship network and topic similarity aspects in the matching mechanisms. They tested researchers preferences regarding existing co-authors vs. new potential candidates with a light-weight user study by asking the participants to rate each recommendation on a scale from one to ten. The findings revealed that participants prefer existing connections with whom collaboration has been already established.

To summarize, the prior work displays a diversity of research and development of people recommender services for scholars with a particular focus on analyzing human-generated content and publishing history. The literature on algorithmic approaches indicates that bibliographical data sets can serve as a valid data source for identifying and recommending social connections. The algorithmic choices encouraged us to approach the area with specific topic modeling methods (TF-IDF) and to analyze the cosine distance³ (Li and Han 2013) between the authors. Although there is apparent interest in creating services for academic partnering, the primary contribution of preceding research lies in the design of new matching mechanisms and algorithms. The evaluation of such systems, thus, is focused on testing the quality criterion of prediction accuracy and little attention is paid to the subjective perceptions of recommendation usefulness and user's intention to follow-up on those. In this article, we experiment on content-based similarity-difference dimensions and specifically focus on the human-centered and subjective evaluation of perceived relevance and related variables.

2.2. Concepts of similarity and diversity

Our approach of diversity-enhancing people recommender systems is founded on the relatively strong consensus that fruitful collaboration and high innovation capability result from complementary viewpoints among a diverse group of actors (Mitchell and Nicholas 2006). Rodan and Galunic (2004) imply that heterogeneous knowledge is of high importance to both overall managerial performance and particularly to innovation performance. However, the factual value of diversity and how it should exactly manifest remain unclear. Despite the extensive literature, the role of both similarity and diversity (as opposite ends of a continuum), particularly in the decision-making of choosing academic collaborators, requires more research, as will be shown in what follows.

The related work discussed in the previous subsection demonstrates a tendency of utilizing similarity-maximizing approaches for recommending content and connections, thus amplifying the effects of homophily bias. The concept of *homophily* has caught the attention of researchers, primarily in social psychology (Lazarsfeld and Merton 1954; Marsden 1987; Moody 2001), as the phenomenon of individuals' natural preference to interact with similar-minded people who share socio-cultural traits. In CSCW and HCI research, homophily has been addressed, for instance, as a predictive and influential factor of online behavior in content preferences (Chang et al. 2014), and audience attraction on social media (Sharma and Cosley 2016). Another vein of research focuses on studying diversity in terms of human, relational and intellectual capital within global organizations to design features that support online communities in collaborative tasks (Muller et al. 2012). Researchers and developers seem to have adopted the similarity-maximizing approach from item recommender systems, using metrics of similarity as the proxy for relevance also in matching peers within organizations (Guy et al. 2010) and scholars in the context of academic collaboration (Heck 2013).

Although homophily might strengthen existing communities, it does not encourage the creation of new ties to further away in the global social network. Some researchers propose that such mechanisms directly lead to the formation of echo chambers that are detrimental to information flow, innovation, and creativity (Jasny et al. 2015; Bessi 2016). Echo chambers have received critique particularly with respect to social media services that divide the user community into camps of different opinions and thus increase polarization in the society (Li et al. 2013; Lee et al. 2014).

At the same time, organizational studies have identified that also diversity can have negative influences on collaborative activities, such as information exchange and decision-making (Graves and Elsass 2005; Hobman et al. 2004). An extensive review (Mannix and Neale 2005) concludes that social differences (i.e., surface-level), such as race and gender, indeed tend to have adverse effects on the ability of groups to function effectively, whereas more profound cognitive dissimilarities, such as differences in expertise or personality, are more often positively related to

team performance. In other words, diversity is strongly linked to the concept of identity, which can make the introduction of diversity challenging in established work cultures.

Following the above mentioned, CSCW research has investigated whether it is possible to overcome the adverse effects of dissimilarities in teams to provoke creativity and productivity. For instance, Dong et al. (2016) found that commitment to a common cause, such as shared goals of the work, bring people together despite cultural differences. Similarly, but from a broader perspective, Ye and Robert Jr. (2017) revealed that collectivism (over individualism) makes people more tolerant to differences in terms of personal values, working styles, skills, and general abilities, thus, embracing individual creativity and work satisfaction. Besides, Rajagopal et al. (2017) investigated how to match peers with dissimilar opinions. The findings demonstrated that matching people with different interpretations of shared interests is more effective in producing positive experiences of breakdown.

Overall, the literature on diversity and homophily contains interesting contradictions, which calls for further empirical research on various forms of similarity or diversity in different types of collaboration. To this end, we seek to uncover how the two concepts are interlinked particularly in the assessment of the relevance of potential scholarly collaborators.

2.3. User-centered evaluation criteria for recommender systems

Historically, research on recommender system has primarily focused on the design of algorithms, underlying the assumption that better algorithms results in better user experience with the systems. Pu et al. (2012) challenges this premise by providing conceptual observation and guidelines on the evaluation criteria for recommender systems. They explicitly emphasize the importance of the user's perception regarding the system qualities. We summarize existing conceptualizations of the recommendation quality as follows: (i) *perceived accuracy* (Pu et al. 2011) – how well recommendations match with users interests defines the trust towards the systems; (ii) *familiarity* (Sinha and Swearingen 2002) – presence of familiar items increase trust towards the system; (iii) *novelty* (Castells et al. 2015) – unexpectedness of received recommendations can affect perceived usefulness of the system; (iv) *diversity* (Nguyen et al. 2014) – receiving diverse items lessens filter bubble thus increasing users' satisfaction and, as a consequence, perceived accuracy of the system.

Knijnenburg et al. (2012) also provide a framework for the user-centered evaluation of recommender systems that extends the system accuracy metric with other relevant measures. For instance, the authors observe correlations between concepts, such as perceived recommendation quality (relevance), choice satisfaction, variety, diversity, effectiveness, and accuracy along with personal characteristics of the user (e.g., trust towards ICT).

To sum up, these types of evaluation criteria focus on subjective user perceptions' in the evaluation of objective aspects of the system. In this article, we do not question the effectiveness of the designed system and its elements, but rather focus on investigating scholars' attitudes towards the recommended people as potential collaborators. The evaluation criteria proposed by the prior research has proven to be effective in the assessment of item recommender systems. In contrast, as objects of recommendation, human individuals contain much more diverse features that influence the evaluation. When assisting people in choosing potential collaborators, the subjective perception of relevance might have various facets and be determined by the need or task for partnering, and, therefore, the metrics should be operationalized accordingly. In this article, we approach relevance criteria with the temporal aspect (i.e., from the perspective of past vs. current research interests), as well as from the perspective of potential collaborative activities with the recommendations.

3. System design

In this section, we first explicate the choice of the data source used for the design of the recommendation algorithm. Next, we outline the data cleaning process and analysis and, finally, describe the user interface developed for the experiment.

3.1. Data source, data cleaning, and analysis

We designed a content-based people recommender system using DBLP, an open bibliographic database of publications records from the majority of Computer Science conferences and journals. The DBLP dataset is a substantial plain ASCII XML file.⁴ The metadata for each record contains more than necessary details for the study and, therefore, requires multiple cleaning procedures.

In the first step, the XML file was parsed using the 'xml.sax'⁵ package targeting on the following tags: article, inproceedings, proceedings, book, incollection, phdthesis, mastersthesis, and www. Then, from the parsed XML file, we extracted 5,847,090 records that consist only of titles, co-authors, publishing years and venues. Next, in the resulted subset we cleaned the titles of publications following three steps: (i) converting letters to lowercase, (ii) removing the English stop words with 'nltk.corpus.stopwords'⁶ function, (iii) removing the digital strings.

The people recommender system runs on the subset of the parsed 5,847,090 records and depends on the input of publication venues of a given user (participant of the study). The detailed data analysis is demonstrated in Figure 1. For a participant, top venues of their publications are given (see step 1 in Figure 1). Then a subset of records is extracted from the parsed DBLP dataset by only those publication venues (step 2). All the titles in the subset records are cleaned as described previously and aggregated to form the corpus profile for each author ever published in those venues (step 3). Those authors who have less than three

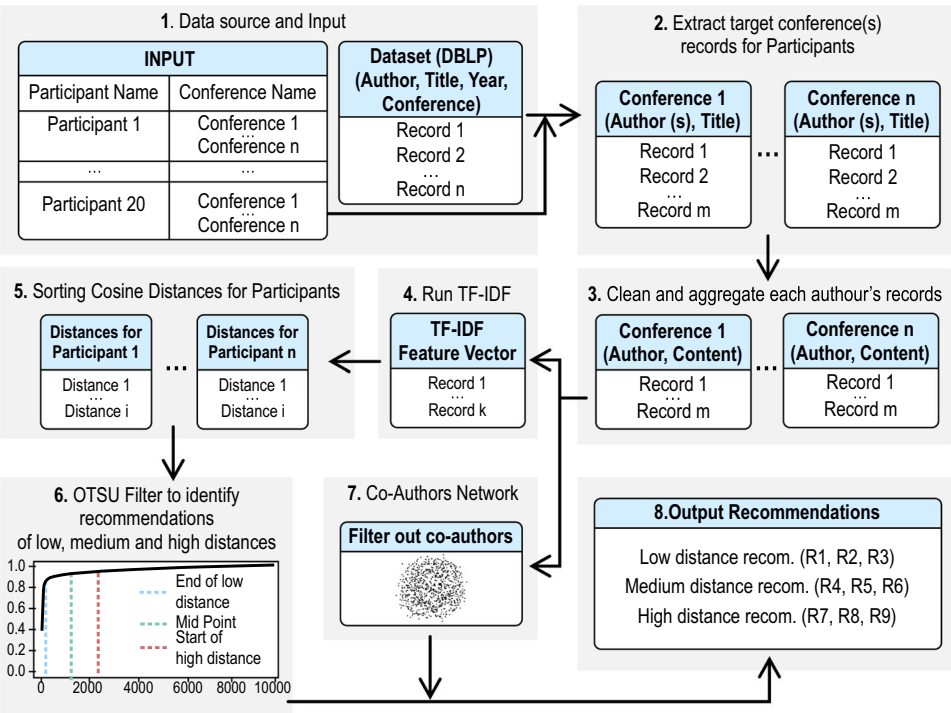


Figure 1. Overview of the DBLP data analysis process to produce people recommendations.

publications in recent five years are filtered out to improve the quality of recommendations. Next, we tokenize the corpus and build the vocabulary with the words that only appear once or show up in more than 95% of the author corpus profiles using ‘CountVectorizer()’ function from scikit-learn.⁷

After corpus tokenization, TF-IDF is applied to the profile model to form feature vectors for each author (step 4). Next, we compute the cosine distances between the given participant and the other authors in the subset records regarding them (step 5). As it is more intuitive to indicate the close distance with a smaller number, we use cosine distance to represent the similarity between two authors. Accordingly, the closest, or the most similar author to a participant will have the smallest cosine distance.

To validate the participants’ preferences on similarity-difference continuum during the user study, we decided to deliver recommendations in the form of three groups of controlled distances – low (high similarity), medium (moderate similar) and high (low similarity). To automatically separate recommendations into such groups, the cosine distances between a participant and the other authors are sorted first, and then the OTSU filter (Otsu 1979) is employed to detect the boundaries between each group (step 6). The OTSU filter calculates the optimum threshold separating the two groups so that their intra-class variance is minimal. As the distribution of the cosine distances follows the power law, we implement

the OTSU filter twice. In the first round, we apply it to the whole sorted cosine distances to detect the boundary between the low distance group and the rest. In the second round, we apply the OTSU filter on the rest without the low distance group to divide medium and high distances groups. Finally, three recommendations that have no co-authorship with the given participant (step 7) and published in the same venues are picked from each distance group as the final output, thus delivering nine recommendations in total (step 8).

3.2. User interface

A single-page web application was deployed in Firebase development platform.⁸ Figure 2 illustrates the User Interface (UI) view with personalized recommendations in the form of a carousel-based list. The UI visualizes all information about authors, which DBLP data set allows to extract: full name, research topics, the list of co-authors, and recent publications. Each section of the UI is expandable if there is additional content available. The publication list represents only works from conferences where both the recommended person and the participant of the experiment have published in. Accordingly, the list of co-authors is taken from those publications only. The topics were generated through bigram analysis on the corpus profiles of each recommended person. We first generate the bigram word pairs using NLTK ‘bigram’ function on all the corpus. Next, we use ‘nlk.ConditionalFreqDist’ to calculate the occurrence of other words by giving a certain word in the corpus. For example, in a bigram word pairs for a word ‘social,’ the word ‘media’ may appear 20 times, while the word ‘compute’ may appear zero times. Then, to generate the authors’ topics, the bigram word pairs are

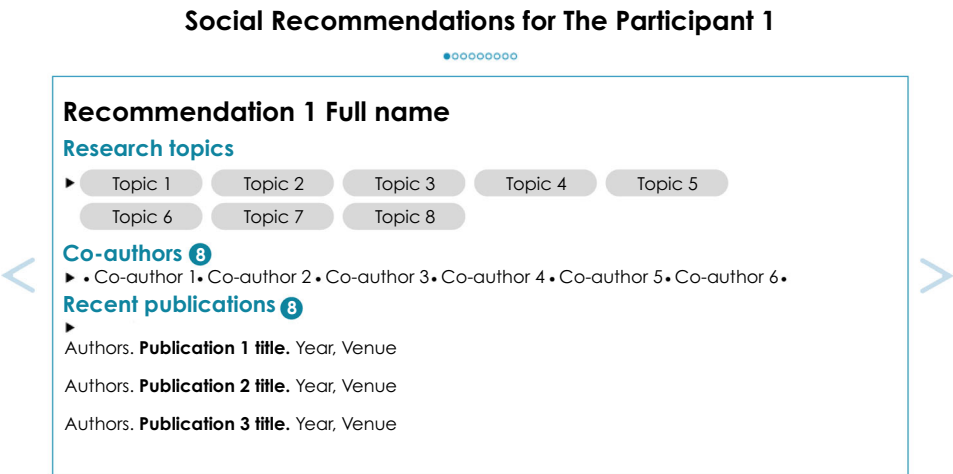


Figure 2. The UI presented to the participant (here with anonymized recommendation).

created from their corpus. After that, we check the conditional frequency of the second word regarding the first word in each pair. If the frequency is equal to or higher than ten (10), we pick this bigram word pair as one of the authors' topics.

4. User study

We designed a user study combining a controlled experiment and a semi-structured interview. By providing participants with real recommendations, we aimed to help them to form their opinion regarding experiment variables. Following the homophily bias, we hypothesize that the lower the cosine distance between the participant and the recommended person (i.e., the similarity of publishing history), the more relevant and similar the recommendation would be perceived.

4.1. Experimental design

In the experiment, the computed cosine distance (content-based distance) is the independent variable, represented as three groups of fellow academics – those with low, medium or high distances. Thus, recommendations of other researchers with *high similarity (low distance)*, *moderate similarity (medium distance)* and *low similarity (high distance)*, with three recommendations from each group were presented to the participants. The participants were not informed of the three groups to avoid biased evaluation, and the presentation order of the altogether nine recommendations was randomized. The evaluation inquired the participants' perceptions about the following dependent variables: relevance, similarity, familiarity and willingness to interact.

4.2. Recruitment and participants

For the experiment, we recruited 18 English-speaking senior researchers who work at two university campuses in Tampere, Finland. Following the assumption that senior researchers often have more needs for finding collaborators, we limited our scope to postdoctoral researchers, professors, or otherwise senior academic positions. For the recruitment, we utilized various e-mail lists to reach relevant faculties, departments, and research groups. In addition to offering the participants a movie ticket for their participation, the recommendations of potential collaborators were also marketed as incentives to take part in the study.

Overall, we had 13 male and five female participants, all based in either of the two universities in the same city. Fourteen of them are Finnish, two Russians, one British, and one Romanian. The ages vary from 32 to 66 (Median: 42, Mean: 45). Seven of the respondents reported their current occupation as Senior Researchers, six as Postdoctoral Researchers, four as Full Professors, and one as an Associate Professor. The most frequent research interest of the participants included

human-computer interaction (10), gaze technologies and interactions (7), wireless technologies (6), interaction design and techniques (6), interfaces and information systems (5), usability/user experience and user-centered design (5), telecommunications and networking (5), virtual reality (3), wellness/health technologies (3). Their academic experience varied from 10 to 46 years (Median: 19, Mean: 20.3).

Figure 3 illustrates the participants' backgrounds and attitudes concerning technology orientation, social openness, activity in networking, and breadth of research interests. Along with the other background information, the figure implies that the respondents represent what we would consider as typical computer science scholars, being technically oriented and curious about research, while displaying variety in their networking practices and interests.

4.3. Procedure and data gathering

The data gathering is comprised of three parts: (i) screening of suitable participants based on their professional position and publishing history before the experiment session. These data were used to prepare personalized recommendations for each participant. (ii) In the experiment session, the participant signed a consent form by filling out an online survey, including also a background questionnaire and numerical evaluations of each recommendation. (iii) The experiment was followed by a semi-structured interview to gather qualitative data about the participants' choices and needs for collaboration. The whole study session lasted from 40 minutes to 1.5 hours, depending on the time the participant

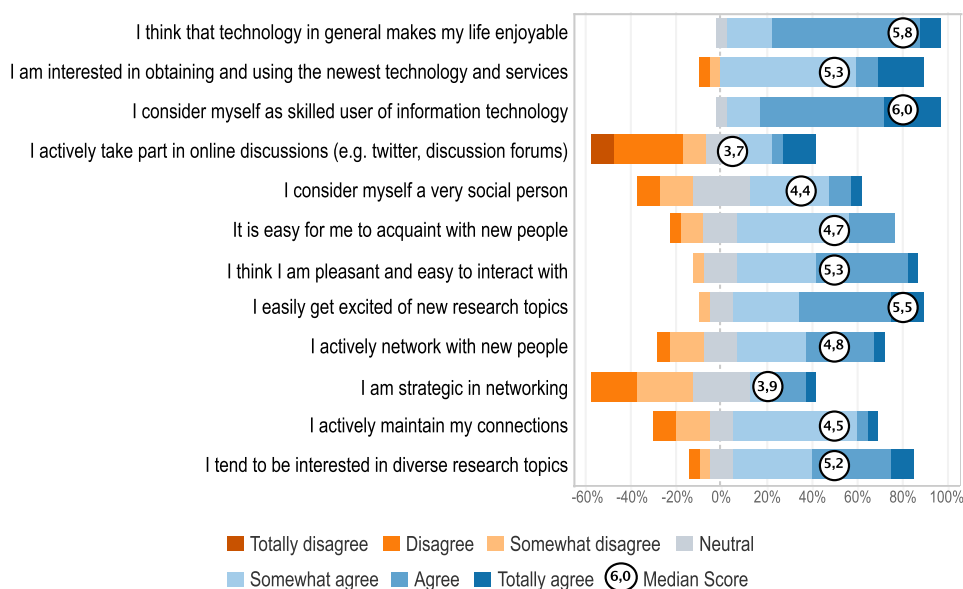


Figure 3. The overview the participants' backgrounds and attitudes.

took to get familiar with and assess the recommendations and how opinionated and expressive they were in the interviewing part. All sessions were audio recorded with the participants' permissions.

Before starting the numerical evaluation, participants were given time to explore all the recommendations and get a general overview of the alternatives. The evaluation was constructed according to four variables (see questionnaire verbatim in Table 1): (i) **perceived relevance** from the perspective of current and past research interests (Q1 and Q2), (ii) expected **willingness to interact** with a recommended person in the context of a scientific conference, including six traditional collaborative activities (Q3-Q8), (iii) levels of **perceived familiarity** – whether or not the user is familiar with the target person, with their research, or with their co-authors (Q9-Q11), (iv) **perceived similarity** between the participant and a recommended person (Q12). The variables were operationalized based on the authors' personal experiences and qualitative research insights on academic collaboration and user experience evaluation. Originally over 20 candidate items were assessed within the project team and with collaborators in an iterative fashion, resulting in the included 12 items. After providing the ratings, participants were asked to explain the scores and their reasoning behind them verbally. The interview questions that are also presented in Table 1 were designed to obtain participants' rationale behind the scoring of recommendations as well as to reveal needs and factors that affect decision-making in academic networking practices.

4.4. Data analysis

Tableau⁹ was used for analysis and visualization of participants' background information, and RStudio¹⁰ for statistical analysis and visualizing the scores in multiple box plots. The experiment has a thrice-repeated within-subjects design with nine categorical data points per participants. We utilized non-parametric Friedman test (Sheldon et al. 1996) and post-hoc analysis with 'Agricolae' package¹¹ in RStudio to identify a statistically significant difference between the participants' ratings of the three recommendations' groups in all questions. To avoid pseudo-replication, we calculated medians of scores given to each group of recommendations. Thus, the input data for the Friedman test consisted of Participant ID, Similarity distance groups (Low, Medium, High) as factors, and medians of scores as values.

As for the qualitative data, the audio recordings from each session were transcribed and resulted in a text file for each participant (with Min 211, Max 1,373 and Median 690 words). The coding procedure consisted of two cycles including elemental, axial and focused methods (Saldaña 2015). At the first cycle, we applied structural coding that allowed us to group data under top-level categories from interview questionnaire (see Table 1): overall impression, collaboration needs, comments about recommended people and their content, essential factors

Table 1. Online survey structure and questions verbatim.

Recommendations evaluation (Likert 7 point scale, 1: completely disagree, 7: completely agree)
(Perceived Relevance) I consider this person is relevant to me from the perspective of: Q1. My current research interests Q2. My past research interests
(Willingness to interact) I consider this person is relevant to me for the following activities at the context of a conference Q3. Asking for advice Q4. Giving advice Q5. Sharing research results and/or ideas Q6. Exploring joint research interests Q7. Spending time at a conference together Q8. Organizing a research visit Which other activities would this person be interesting to interact with?
(Perceived Familiarity) To what extent are you familiar with this person? Q9. I am familiar with the research of this person Q10. I know by name some of the coauthors of this person Q11. I know him/her in person
(Perceived Similarity) Q12. How similar to you do you find this person’s research areas? Very different (1) - Very similar (7)
Interview questions
1. Please tell about your overall impression about the recommended people. 2. Could you please specify what kind of needs for collaboration you normally have? 3. Choose one the most relevant and irrelevant persons and briefly explain why you gave them the scores you did. 4. About whom was it particularly hard to make an evaluation? 5. Which of the presented information about the recommended person do you consider relevant/meaningful in professional matching? 6. What would other factors about the people be important when considering with whom to collaborate? 7. Overall, how would you feel if a system like this gave you recommendations of whom to collaborate with? 8. Would you prefer to receive automatic notifications during a professional event or look for relevant people by yourself?

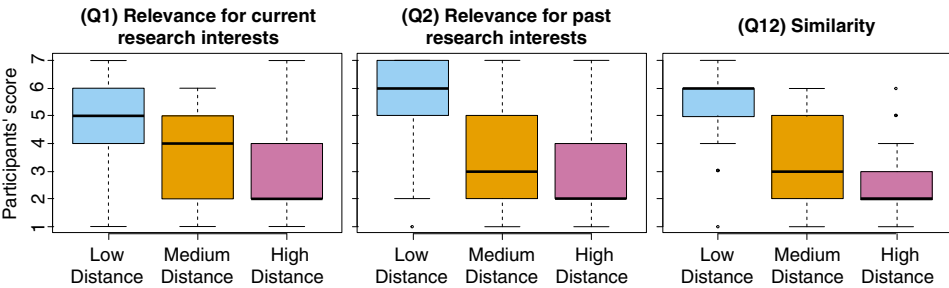


Figure 4. The distribution of participants’ scores regarding perceived relevance and similarity.

in social matching, attitudes towards ICT-mediated professional social matching. Then, for the data in each category, we utilized line-by-line analysis and deconstruction of data into emerging categories, which were further reconstituted resulting in subcategories, linkages, and relationships. Finally, the focused coding was applied to identify the most frequent codes and organize them into emerging themes.

5. Results

We first report the quantitative and qualitative results on perceived relevance, similarity, familiarity and willingness to interact. Then, we discuss the needs and important factors in research collaboration, which might have affected the participants’ decision-making on potential social interactions.

5.1. Perceived relevance and similarity

5.1.1. Quantitative findings

The participants evaluated the perceived relevance of given recommendations from two perspectives (see Figure 4): relevance for current (Q1) and past research interests (Q2). To summarize, the scores were found to be **consistent** with the computer-defined cosine distance. The highest scores were given to the group of recommendations with low distance, while those with the high distance generally received the lowest grades. This indicates the prevalence of homophily bias (preferring most similar researchers) in this sample of participants. Nevertheless, the data also reveals high ratings of perceived relevance for the group of recommendations with low similarity. These matches have received scores of five and higher (appears 16 times) that indicates some participants’ interest in dissimilarity and openness towards new opportunities.

A Friedman test and post-hoc analysis indicated a statistically significant difference in the scores of each group of recommendations (see Figure 5). Ratings of relevance for the past and current research interests demonstrated almost equal

Question	Test Statistics	Distance group	Ranks	Friedman's Post Hoc Analysis	
				Low Distance	Medium Distance
Relevance for current research interests (Q1)	N=18, Chi-Square=16.5 DF = 2, p<.001***	Low Distance	47.5		
		Medium Distance	35	<.01**	
		High Distance	25.5	<.001***	<.05*
Relevance fo past research interests (Q2)	N=18, Chi-Square=24.03 DF = 2, p<.001***	Low Distance	50		
		Medium Distance	34.5	<.001***	
		High Distance	23.5	<.001***	<.01**
Similarity (Q12)	N=18, Chi-Square=25.21 DF = 2, p<.001***	Low Distance	52		
		Medium Distance	32	<.001***	
		High Distance	24	<.001***	<.05*

Figure 5. Results of Friedman's Test and post-hoc analysis. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

results, meaning that participants were consistent in their scores independently of the temporal perspective.

The quantitative results of perceived similarity (Q12) also demonstrate the tendency of a significant difference in given scores (See Figure 4). Accordingly, matches with low distance are graded as most similar, matches with medium distance as somewhat similar and those with high distance – least similar. This consolidates validity of cosine distances and OTSU filter as a method for identifying thresholds of three degrees of similarity. Friedman test results consolidate a significant difference between recommendations groups (See Figure 5).

5.1.2. Qualitative findings

The verbal feedback about relevance and similarity is generally in line with the quantitative results. To provide an overview of participants' comments, we collected illustrative examples in Table 2, sorted according to the three groups of similarity distance. Although the participants were unaware of the three different similarity distances, in their feedback they distinguish between different degrees of perceived relevance by using phrases like '*very/most relevant*,' '*somewhat relevant/not exact match*' and '*irrelevant/totally irrelevant*.'

Feedback about the outliers (in Q1 – 9 cases, in Q2 – 7 cases) – recommendations of high distance rated as relevant – has revealed that such recommendations relate to participants' current research interests with potential for future directions, or because of surprising topics appeared in their profiles. Such cases hint about interest in dissimilarity and openness for new opportunities, as quotes 1 and 2 illustrate.

1)“(R8) This person has really interesting research topics. [...] I thought it would be fun to get to know something else, something different from my own research focus. I would definitely go to see their poster. Maybe it would be interesting for my future research. This is an interesting and fun recommendation!” (P1, 50 y.o., Finnish female, Post-doc)

Table 2. Examples of participants’ verbal feedback about given recommendations.

Low distance (high similarity)	Medium distance (moderate similarity)	High distance (low similarity)
<p>(R2) I am familiar with this person and his coauthors but not in person because I didn’t collaborate with this Greek team directly. However, I was reviewing their publications, particularly, the most recent paper on the list. I consider this recommendation is the most relevant to me. I would surely like to meet in person. (P3, 32 y.o., Russian male, Senior Researcher)</p>	<p>(R4) I guess this person is some research leader or professor. Some papers are close to what I am doing, but not an exact match. I would probably find a lot of interesting discussion with him. He seems to be more experienced than me. There is a clear shared interest. I do not remember hearing about him and do not recognize any of co-authors. (P8, 33 y.o., Finnish male, Postdoctoral Researcher)</p>	<p>(R7) I have never seen this person before. There are thousands of young researchers, I am not surprised. I have no idea why the system recommends me this person because his topics are totally different from mine. I do not know any of co-authors and topics do not seem to be relevant. (P7, 56 y.o., Finnish male, Senior Researcher)</p>
<p>P3 research interests: wireless technologies, telecommunication and networking</p>	<p>P8 research interests: Internet-of-Things, wireless technology, radio networks, and positioning</p>	<p>P7 research interests: Human-computer interaction, VR, gaze technologies and interactions</p>
<p>R2 research topics: wireless technology, cellular systems, sensor networks</p>	<p>R4 research topics: Neural networks, stochastic decoding</p>	<p>R7 research topics: Crowd science, human-computer interaction</p>

2)“(R9) This person is from a different field, and the research has interesting aspects. The last paper in the list is the most interesting: it is about something similar we have been doing recently but not published yet! So, I have to check it and maybe contact the authors.” (P2, 42 y.o, Finnish male, Senior Researcher)

Some participants also mentioned that their first impression about recommended people had changed when they started to check the recommendation’s profiles in details (see quote 3).

3) “That was a good idea that you gave me to check all recommendations first because initial reaction was different, but when I started thinking and realized that my first impression maybe was not correct. When you start thinking about recommendations’ relevance further, there might be some changes. So some of them are not that irrelevant as I thought at first.” (P16, 50 y.o., Finnish female, Senior Researcher)

Hence, decision-making on perceived relevance was found to be influenced by the participants’ estimation about how topics of recommended people match with their own – whether they are similar, very different or complementary. In

this regard, few participants addressed that evaluating the similarity between them and the recommended people is a challenging task. First, they mentioned employing different scales of comparison (5 cases) – assessing a recommended person from the perspective of all research fields in the world or specific focus areas. The second factor refers to the temporal aspect (5 cases): as interests and research directions tend to change, the evaluation of the similarity depends on the chosen time frame. Participants also pointed out that the estimation of relevance and optimal level of similarity for specific collaboration are context dependent (4 cases): some tasks would require cooperation with diverse people, while other tasks benefit from similarity. Thus, when looking for candidates with a distinct set of skills, people should also define shared interests or goals to make a prospective collaboration fruitful. Participants also acknowledged the possible adverse effects of receiving recommendations of people who are very similar (4 cases). For example, in a professional context, high similarity of interests might result in competition, and social interaction with such people will require choosing specific communication strategies.

5.2. Familiarity

5.2.1. Quantitative findings

Familiarity variable was evaluated from three perspectives: familiarity with the research topics (Q9); recognizing some of the co-authors' names (Q10); and knowing the match in person (Q11). The Figure 6 depicts that the scores of recommendations with a low distance in all questions are widely distributed, while groups of the medium and high distances received mostly negative ratings. According to scores distribution, participants are mainly familiar with the research of recommendations and aware of many co-authors of recommended people in low distance group. Besides, there are only a few people whom participants know in person. The results of the Friedman test yield a statistically significant difference in given scores (see Figure 7). The post-hoc analysis reveals significant variance only in groups of low vs. medium and low vs. high distances.

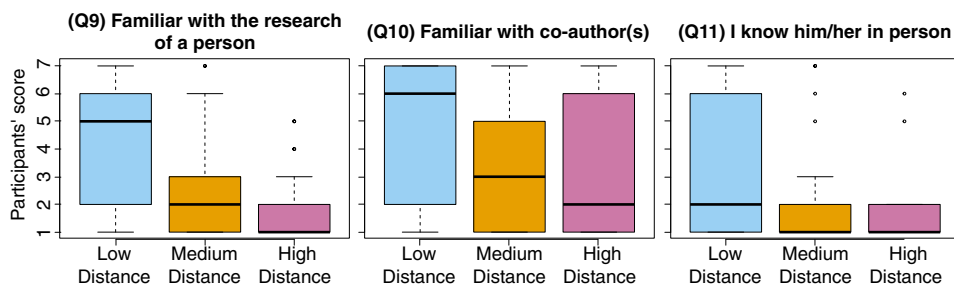


Figure 6. The distribution of participants scores regarding perceived familiarity.

Question	Test Statistics	Distance group	Ranks	Friedman's Post Hoc Analysis	
				Low Distance	Medium Distance
Familiar with the research of a person (Q9)	N=18, Chi-Square=21.26 DF = 2, p<.001***	Low Distance	48		
		Medium Distance	32.5	<.001***	
		High Distance	27.5	<.001***	.1108
Familiar with co-author(s) (Q10)	N=18, Chi-Square=12.04 DF = 2, p<.01**	Low Distance	46		
		Medium Distance	33.5	<.01**	
		High Distance	28.5	<.001***	.2597
Know match in person (Q11)	N=18, Chi-Square=23.4 DF = 2, p<.001***	Low Distance	48		
		Medium Distance	30.5	<.001***	
		High Distance	29.5	<.001***	.7048

Figure 7. Results of Friedman’s Test and post-hoc analysis. Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

5.2.2. Qualitative findings

In general, participants were positively surprised to see many unfamiliar people in the list of recommendations and were happy to increase their awareness about new researchers in related scientific fields (see quote 4 and 5). At the same time, they appreciated receiving suggestions about people with established ties or well-known figures and interpreted it as confirmation of the algorithm validity.

- 4) I think the 1st one was most familiar, he is a good match. For the others, I started to wonder about their research topics. I tried to see co-authors identifying relation to me and then to topics. I think it is good that the majority of recommendations are unfamiliar because it is a people recommender system. (P6, 35 y.o., Finnish male, Postdoctoral Researcher)
- 5) I was curious about why I do not know some of the people. I was trying to analyze, and I think there might be a time delay in our academic activities or maybe there is a gap regarding research communities. (P3, 32 y.o., Russian male, Senior Researcher)

In the interview, participants addressed that already known people are less exciting recommendations (quotes 6 and 7). Even though we intentionally filtered out all the co-authors, the bibliographic data prevents understanding of the actual social relationships between researchers. Thus, in some cases (9 recommendations out of 54), the system recommended people from very close social circles, for instance, peers from academic projects or colleagues with whom one had not co-authored publications but interact daily (quote 8).

- 6) “(R3) It seems like I know him. His areas of research match with mine, I would say, by 90%. I know some of his co-authors even in person. This is the most interesting recommendation but not surprising, because I know him and even know his Ph.D. students.” (P5, 42 y.o., Finnish male, Full Professor)
- 7) “The first and last person in the list I already know. They are obviously relevant recommendations. But the rest is more interesting because I do not know them.” (P10, 40 y.o., Finnish male, Senior Researcher)

8) “(R2) One interesting observation is that one of the recommended people is my roommate. We have similar topics and plenty of other commonalities. We are cooperating mostly by talking. I think, we have never written papers together, but we have shared co-authors.” (P4, 56 y.o., Russian male, Senior Researcher).

5.3. Willingness to interact

5.3.1. Quantitative findings

Evaluation of the willingness to interact with recommended people reflects six predefined scenarios of face-to-face interaction or follow-up collaboration at the context of a conference (see Figure 8): (Q3) asking advice, (Q4) giving advice, (Q5) sharing research ideas, (Q6) exploring joint research topics, (Q7) spending time together, and (Q8) organizing a research visit. The distribution of scores demonstrates that participants seemed to have a very positive attitude towards engaging in a low-threshold interaction like sharing research ideas, exploring common research topics and spending time together, particularly with most similar people. Scores given to medium and high distance groups illustrate variance of opinion with a neutral attitude on average. Interestingly, all the six interaction scenarios yielded very similar results, even in the more long-term follow-up action of organizing a research visit. The results of the Friedman test (see Figure 9) depicts a statistically significant difference in given scores only in groups of low vs. high and low vs. medium distance.

5.3.2. Qualitative findings

In general, participants have diversified opinions regarding the estimation of any interactions and follow-up activities. The majority of participants (12 cases)

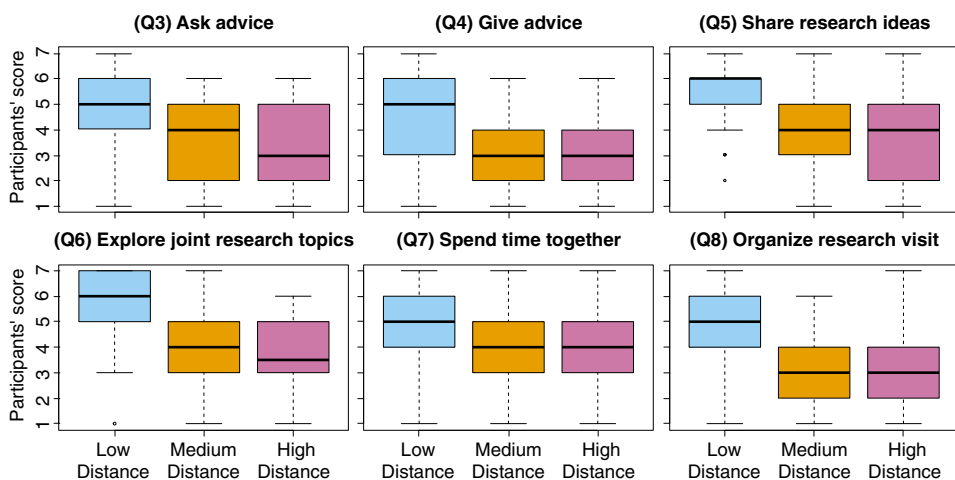


Figure 8. The boxplots visualize participants scores distribution about scenarios of willingness to interact.

Question	Test Statistics	Distance group	Ranks	Friedman's Post Hoc Analysis	
				Low Distance	Medium Distance
Ask advice (Q3)	N=18, Chi-Square=13.54 DF = 2, p<.001***	Low Distance	47		
		Medium Distance	32.5	<.01**	
		High Distance	28.5	<.001***	.3589
Give advice (Q4)	N=18, Chi-Square=9.17 DF = 2, p<.01**	Low Distance	45		
		Medium Distance	31.5	<.01**	
		High Distance	31.5	<.01**	1
Share research ideas (Q5)	N=18, Chi-Square=18.453 DF = 2, p<.001***	Low Distance	48.5		
		Medium Distance	32	<.001***	
		High Distance	27.5	<.001***	.2321
Explore joint research topics (Q6)	N=18, Chi-Square=20.818 DF = 2, p<.001***	Low Distance	50.5		
		Medium Distance	32.5	<.001***	
		High Distance	25	<.001***	.059
Spend time together (Q7)	N=18, Chi-Square=16.754 DF = 2, p<.001***	Low Distance	49		
		Medium Distance	30.5	<.001***	
		High Distance	28.5	<.001***	.633
Organize research visit (Q8)	N=18, Chi-Square=22.286 DF = 2, p<.001***	Low Distance	50		
		Medium Distance	32	<.001***	
		High Distance	26	<.001***	.0831

Figure 9. Results of Friedman’s Test and post-hoc analysis. Significance codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

consider such social activities to happen naturally at conferences and do not necessarily require high investments of time in collaboration. For some, it was hard to envision willingness to interact with unfamiliar people based on papers titles and topics of the research. Thus, some participants (8 cases) admitted that they would like to learn more about a recommended person before taking any decisions on social interaction. Some emphasize (5 cases) that even in a real context of visiting the conference, it might be challenging to find relevant people in a crowd and contextualized use of such recommender systems might simplify the process and encourage social interaction with unfamiliar people (see quote 9).

9) “Such system should narrow the focus to serve specific purposes [...] I think about a scenario where I am going to the conference, so then I can define ‘show me people who are relevant to this event’ and it would give me a sense of community around it [...] After all these years you sometimes stand somewhere in the corner of a conference hall not knowing anybody. Of course, I can start communicating with random people, but it would be much more efficient if the system can suggest already somewhat relevant people and provide with tickets to talk.” (P11, 53 y.o., Finnish female, Full professor)

5.4. Needs and important factors in research collaboration

When specifying crucial needs for collaboration, participants mentioned different activities. The most frequent reasons are understandable when considering

senior academics: seeking academic and industrial partners for funding applications (appears in 12 answers) and knowledge sharing as the way of indirect cooperation (12 answers, e.g., exchange of data or finding relevant publications in the topic of interest). Some participants also pointed out a need for people without conflicts of interest (5 answers) such as pre-examiners, reviewers, editors or opponents, who are highly demanded and complicated to find. Another reason was research mobility (5 answers), which calls for cooperation with particularly international universities or companies.

Interestingly, whereas similarity was generally considered to be a significant aspect, the above-mentioned collaborative relationships demand heterogeneity of methodological skills, research areas, or social networks. The participants also emphasized that needs for collaboration are occasional and it will be useful to contextualize the recommender system to specific scenarios, for instance, make it particularly location- and event-based or expert-finder. In their opinion, this will ensure the reasoning for using a service and motivate to follow-up on recommendations (see quote 10).

10) “I would be interested in such a system to explore people who are visiting the same conference in advance and filter them based on similarity or relevance. [...] It will help me to revise recommendations faster. Let’s say for the event-based mobile application it will be great to inform me when a person visits the event and recommend me to meet him there. If a notification to interact comes in the middle of a street, then I doubt it will work. However, if it will happen at the conference I, of course, will try to follow-up on recommendations.” (P8, 33 y.o., Finnish male, Postdoctoral Researcher)

The participants also specified factors that matter to them when seeking professional collaboration. First, the majority (14 replies) addressed the importance of affiliation and the current position of candidates. From their perspective, it can tell a lot about the seniority, availability, and potential interest of the people. Besides, considering the relatively high migration of researchers to non-academic positions, it can indicate whether potential cooperators are still pursuing an academic career. Furthermore, many factors can be implicitly obtained from publications. For instance, the quantity, citation rates, and quality of papers might reveal information about the maturity of a researcher, their topics of interests as well as information about their community. For many (8 replies), these aspects play a significant role when aiming to approach unfamiliar scientists (see quote 11).

11) “There are different influence groups with leading experts which are often competitors. So, based on co-authors of a match in his publication lists I can instantly interpret that he belongs to particular influence group. That can help in decision-making whether to collaborate or not and carefully choose the communication strategy.” (P5, 42 y.o., Finnish male, Full Professor)

Discussion on seniority level brought out various opinions (7 replies). In general, seniority plays a considerable role: for instance, in tasks that require straightforward ability to make decisions (e.g., project planning) it is essential to be in contact with mature researchers, while some practical implementations could be performed in cooperation with students (e.g., assisting a course). Other scenarios might call for open-mindedness regarding this aspect, like in the following example:

12) "The seniority level does not matter to me that much. Sometimes junior people are more creative and innovative. [...] So we should never think about seniority levels. More senior people might have much information, but at the same time too narrow in their vision and interests. Of course, it depends: for consultancy, I might prefer to contact senior people, while for generating new ideas and brainstorming I will be more interested in collaborating with young researchers." (P15, 50 y.o., Finnish female, Senior Researcher)

Additionally, participants address the personal chemistry factor (7 replies), specifically for cases of direct collaboration it might be crucial regarding the efficiency of interpersonal relationships and teamwork (see quote 13):

(13) Chemistry plays a significant role – we need some basis for communication. It should be a person with whom it is nice to sit talk and drink coffee in addition to work practicalities. (P15, 65 y.o., British male, Senior Researcher)

Thus, participants highlighted that this factor is highly demanded yet unfeasible to be integrated into the system because in their opinion personality compatibility can be assessed only after continuous interactions.

6. Discussion

While prior research has aimed at creating meaningful professional connections with the help of people recommender systems, little attention has been put on evaluating the subjective perceptions of the recommendation relevance. We emphasize that recommending is different from predicting new connections (McNee et al. 2006): to design services that can meaningfully enhance professional collaboration, algorithms should go beyond reproducing or strengthening the typical human bias.

In the following, we first summarize our findings and reflect on their novelty and relevance. Next, we provide a discussion on limitations and future work.

6.1. Summary of the results

We presented the results of an experiment on computer science researchers' preferences regarding potential collaborators of different similarity levels with a DBLP-based recommender system. With 18 senior scholars in areas related to CS,

we tested how the dependent factors of perceived relevance, similarity, familiarity and willingness to interact are related to the independent variable of objective similarity measurement in terms of publication history.

The findings reveal that (1) the homophily bias is evident also in scholars' intuitive assessments of relevance and willingness to interact, and (2) there is a mismatch between people's intuitive choices and the deliberate intentions in decision-making on potential collaborators.

Considering our first research question about which level of system-defined similarity is preferred in participants' evaluations, the findings demonstrate the highest ratings for most similar people. Methodologically, the subjective evaluations of different similarity levels seem to consolidate the system design, particularly the efficiency of the OTSU filter in identifying different levels on the similarity–difference continuum. In other words, even the relatively simple analytics procedure with scarce data seemed to work sufficiently, and the publication data represented the participants' topics accurately enough. While the norm in such data analytics tends to stress the need for Big Data (e.g., Hoang et al. 2017), it appears that for people recommendations the systems could suffice with rather simple datasets as long as the recommender engine logic is well designed. The findings imply that the participants were able to retrieve useful suggestions and, for the majority, the evaluation process with the operationalized variables was straightforward.

Regarding the second research question about academics' needs and expectations in professional collaboration, the results demonstrate that the optimal area on the similarity-difference continuum highly depends on the type and context of collaboration. For instance, crucial factors in direct cooperation, such as personal compatibility and similarity of attitudes and beliefs, are not as emphasized in short-term and indirect professional interactions (e.g., consultancy type of cooperation) as in long-term collaboration. Furthermore, the nature of the collaboration task might influence the perceived relevance of potential candidates, for example, regarding the complementarity of professional roles, skills, and knowledge.

As a methodological contribution for studying user perceptions, we operationalized the concepts of perceived relevance, similarity, familiarity and willingness to interact as subjective evaluation measures for the context of academic collaboration. This helps to uncover some of the experiential aspects of these concepts and quantitatively assess how individuals consider people recommendations. To complement the reductionist measures, the qualitative findings reveal more complex and nuanced aspects that should be addressed in the design and evaluation of people recommendations for professional partnering.

Following the participants' rationale about important factors in collaboration, we propose that the diversity of recommendations in professional social matching could be enhanced through several dimensions or criteria of relevance:

- *similarity* in terms of background, attitudes, values, beliefs, goals and intentions (e.g., research aims). Previous research addressed that similarity of such qualities can raise cohesion or so-called ‘affinity’ (Moreland and Zajonc 1982) in interpersonal relationships, and can even reduce adverse effects of individual dissimilarities (Dong et al. 2016) in collaborative work;
- *complementarity* in terms of professional roles, skills, knowledge, and social capital. In this context, complementarity is beyond pure diversity, as discussed by Mitchell and Nicholas (2006). It should enable relevant opportunities for collaboration by identifying beneficial intersections between individuals’ qualities;
- *compatibility* for direct cooperation in terms of being mentally, socially, morally or emotionally close to each other. This aspect was partially emphasized by Bozeman et al. (2013), who define collaboration as a process of knowledge production, in which compatibility of such qualities can establish trustful, joyful and personally valued cooperation;
- *approachability/logistics* – the availability of a person for direct or indirect interaction in terms of physical proximity as well as social and organizational distance. This dimension echoes with ‘collaboration readiness’ conceptualized by Olson and Olson (2000), who calls for better technological solutions to enable smooth communication and interaction practices for distributed collaboration.

6.2. Limitations and future work

We selected a mixed-method approach to enable a broad understanding of our research questions. By combining a controlled experiment with qualitative face-to-face interviews, we intentionally limited the sample size and compromised generalizability with deeper qualitative understanding. In the same vein, the participants represent culturally the same geographical area, which means that the generalizability of the findings to the general population of scholars is limited.

Nevertheless, our method allowed us to observe the decision-making process on collaboration in the actual context of using a people recommender system. Additionally, it helped to engage participants in the discussion on potentially relevant partners concerning similarity, complementarity, and other essential qualities or factors when seeking collaboration. We could not have elicited some of the interview findings without having the task of evaluating the recommendations in situ. In fact, our prior research experience suggests that qualitative exploration of human needs, wishes, and expectations often benefits from providing a design artifact that can help a participant to form opinions on abstract concepts and speculated behavior.

6.2.1. Data set limitations

The limitations on data set might explain some of the mismatches in participants' quantitative and qualitative feedback on recommendations. A central limitation of DBLP is that only the publication titles (without abstracts) are available, which limits the content analysis and compromises the accuracy of the user profiles. For example, some participants addressed that it was problematic to assess the relevance of junior researchers who had a small number of papers (Min 3 in the reported study). Limiting the data sample to only those who have extensive publication history would be advantageous for the accuracy of topic modeling and providing comprehensive pictures of the assessed individuals. However, we intentionally wanted to introduce both junior and senior researchers and reveal how they could be appreciated in different contexts of cooperation, as well as evaluate the role of seniority as a possibly influential factor in collaboration processes.

Another limitation of the DBLP data set is that it provides limited understanding to the social ties between researchers. Even though we excluded all the co-authors of each participant from the recommendations, some were recommended people with whom they occasionally interact (9 cases out of total 54 recommendations). The proportion of very familiar people was relatively small, so this can be argued to have an relatively small effect on the perceptions of the system validity. By applying alternative data sources, if practically feasible, it would be possible to implement more advanced analysis of social networks and, thus, prevent recommending already known people.

6.2.2. Homophily bias

Even though the participants of the experiment were unaware of the similarity distance groups, the quantitative ratings of recommendations indeed demonstrated the tendency of researchers preferring most similar people. Only a few participants assessed recommendations with high distances as exciting, surprising and worthy of exploration for potential follow-up. At the same time, the qualitative feedback provides evidence of aspects in a research collaboration that require access to both similar and different others. In the following, we discuss two possible reasons behind the apparent homophily bias, also related to the methodological validity of the study:

(i) *The evaluation was largely based on first impression.* Studies on the cognitive processes of choice (Kahneman 2003; Stanovich and West 2000) distinguish between two modes influencing humans' decision-making – so-called 'system 1' (effortless, intuition-based judgment) and 'system 2' (rational and reasoning-based). In our experimental setup, it seems that most of the participants were primarily relying on their intuition and did not engage in more rational or reflective reasoning in their evaluation. Therefore, the first impression about recommendations was likely driven by the homophily bias, thus explaining the numerical evaluations. At the same time, after rationalizing the matter during the

interview, and reflecting with specific practical collaboration scenarios, the participants started appreciating different types of diversity between themselves and the evaluated person. As for considerations for design, this finding calls for user interfaces that support reflection and multi-dimensional analysis of the collaboration potential with a given recommendation. The current norm in recommender systems and leisurely social matching is based on *hastiness*: using simplistic profiles and simple mechanisms for selecting or discarding recommendations. We argue that such UI and interaction mechanisms do not fit with the goal of identifying optimal academic collaborators.

(ii) *Lacking a timely need for collaboration.* At the moment of the experiment, all of the participants were involved in research projects where the consortium was already built, and they did not report having any urgent needs or requirements for finding new collaborators. Therefore, the estimation of relevance was mostly formed according to their general picture of an ideal collaborator. Nevertheless, the majority emphasized the occasional need to utilize such people recommender systems for research networking and collaboration. This raises two design considerations. First, this calls for user interfaces that support keeping track of different types of collaboration needs in the often so scattered work of academics; most simply, the user could be reminded about their different professional activities upon receiving a recommendation. Second, this calls for context awareness (Mayer et al. 2015b) in timing the recommendation. For example, rather than having separate services for professional matching, the recommendations could be tied to services that academics typically use to seek for suitable collaborators.

7. Conclusions

We evaluated scholars' perceptions of relevance about potential collaborators representing different levels of similarity, utilizing a bibliography-based people recommender system. We operationalized the concept of perceived relevance, familiarity, similarity and willingness to interact within the context of evaluating prospective collaboration. By showing how these variables match with system-defined similarity in bibliography data, we revealed the asymmetry of scholars' intuition-based evaluation and their intentions. The quantitative results demonstrated the effects of homophily bias (preference of most similar others) to perceived relevance, while qualitative findings identify important factors for collaboration that naturally require connection with people of complementary expertise. Compared to the evaluation methods used in item recommenders, the findings demonstrate that people recommender systems require more advanced models and logics that go beyond predicting ties or optimizing for accuracy. From cognitive psychology perspective, assessment of potential partnering with recommended people is a complicated task that should not rely only on the

first impression. Considering the long-term and reciprocal nature of professional collaboration, social matching of scholars calls for domain-specific solutions.

Acknowledgments

This work was conducted under Business Finland project Big Match (3166/31/2017 and 3074/31/2017). We owe special gratitude to Tuomas Kaittola for the implementation of the system. We thank all the members of the project who contributed to the design of the experiment.

Notes

- ¹ DBLP database – <http://dblp.org/statistics/recordsindbpl.html>
- ² TF-IDF is a numerical statistic that is intended to reflect the importance of a word to a document in a collection or corpus.
- ³ Cosine distance is a measure of angular similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them.
- ⁴ The DBLP dataset (2.0 G) can be downloaded from the official website – <https://dblp.uni-trier.de/xml/>
- ⁵ XML.SAX package provides a number of modules which implement the Simple API for XML (SAX) interface for Python – <https://docs.python.org/2/library/xml.sax.html>
- ⁶ The Natural Language Toolkit (NLTK) is a Python-based suite of libraries and programs for symbolic and statistical natural language processing. Official NLTK website – <http://www.nltk.org>
- ⁷ Scikit-learn is a tool for data mining and data analysis in Python – <http://scikitlearn.org/stable/about.html>
- ⁸ Firebase is a mobile and web application development platform which operates on Google infrastructure – <https://firebase.google.com/>
- ⁹ Tableau – a data analytics and visualization tool – <https://www.tableau.com>
- ¹⁰ An open-source integrated development environment for R, a programming language for statistical computing and graphics – <https://www.rstudio.com>
- ¹¹ Agricolae documentation – <https://www.rdocumentation.org/packages/agricolae/versions/1.2-8>

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ackerman, Mark S.; and David W. McDonald (1996). Answer Garden 2: merging organizational memory with collaborative help. *CSCW'96. Proceedings of the ACM Conference on Computer Supported Cooperative Work, Boston, Massachusetts, USA, 16–20 November 1996*. New York: ACM, pp. 97–105.
- Adamopoulos, Panagiotis; and Alexander Tuzhilin (2015). On unexpectedness in recommender systems: Or how to better expect the unexpected. *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 4, January 2015, pp. 54.
- Arens-Volland, Andreas; and Yannick Naudet (2016). Personalized recommender system for event attendees. *SMAP'16. International Workshop on Semantic and Social Media Adaptation and Personalization, Thessaloniki, Greece, 20–21 October 2016*. IEEE, pp. 65–70.
- Argote, Linda; and Ron Ophir (2002). Intraorganizational learning. In: J. A. C. Baum (ed.): *The Blackwell Companion to Organizations*. Hoboken, New Jersey: Blackwell Publishers Ltd., Chapt. 8, pp. 181–207.
- Beel, Joeran; Bela Gipp; Stefan Langer; and Corinna Breitingner (2016). Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries*, vol. 17, no. 4, November 2016, pp. 305–338.
- Beham, Günter; Barbara Kump; Tobias Ley; and Stefanie Lindstaedt (2010). Recommending knowledgeable people in a work-integrated learning system. *Procedia Computer Science*, vol. 1, no. 2, September 2010, pp. 2783–2792.
- Bessi, Alessandro (2016). Personality traits and echo chambers on facebook. *Computers in Human Behavior*, vol. 65, December 2016, pp. 319–324.
- Bozeman, Barry; and Elizabeth Corley (2004). Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, vol. 33, no. 4, April 2004, pp. 599–616.
- Bozeman, Barry; Daniel Fay; and Catherine P. Slade (2013). Research collaboration in universities and academic entrepreneurship: the-state-of-the-art. *The Journal of Technology Transfer*, vol. 38, no. 1, February 2013, pp. 1–67.
- Brusilovsky, Peter; Jung Sun Oh; Claudia López; Denis Parra; and Wei Jeng (2017). Linking information and people in a social system for academic conferences. *New Review of Hypermedia and Multimedia*, vol. 23, no. 2, June 2017, pp. 81–111.
- Burt, Ronald S. (2017). Structural holes versus network closure as social capital. In: N. Lin, K. Cook, and R. S. Burt (eds.): *Social capital*. New York: Routledge, Chapt. 2, pp. 31–56.
- Börner, Katy; Noshir Contractor; Holly J Falk-Krzesinski; Stephen M. Fiore; Kara L. Hall; Joann Keyton; Bonnie Spring; Daniel Stokols; William Trochim; and Brian Uzzi (2010). A multi-level systems perspective for the science of team science. *Science Translational Medicine*, vol. 2, no. 49, September 2010, pp. 1–5.
- Castells, Pablo; Neil J. Hurley; and Saul Vargas (2015). Novelty and diversity in recommender systems. In: F. Ricci, L. Rokach, and B. Shapira (eds.): *Recommender Systems Handbook*. Boston, MA: Springer, Chapt. 26, pp. 881–918.
- Chang, Shuo; Vikas Kumar; Eric Gilbert; and Loren G. Terveen (2014). Specialization, homophily, and gender in a social curation site: findings from pinterest. *CSCW'14. Proceedings of the ACM conference on Computer Supported Cooperative Work and Social Computing, Baltimore, Maryland, USA, 15–19 February 2014*. New York: ACM, pp. 674–686.
- Chin, Alvin; Bin Xu; Chen Zhao; and Xia Wang (2014). From offline to online: connecting people with a mobile social networking application at a conference. *Chinese CHI '14. Proceedings of the Second International Symposium of Chinese CHI, Toronto, Ontario, Canada, 26–27 April 2014*. New York: ACM, pp. 40–49.
- Cox, Donna; Volodymyr Kindratenko; and David Pointer (2003). IntelliBadge TM: towards providing location-aware value-added services at academic conferences. In: A. K. Dey, A. Schmidt,

- and J. F. McCarthy (eds.): *UbiComp 2003. International Conference on Ubiquitous Computing, Seattle, WA, USA, 12–15 October 2003*, Vol. 2864. Berlin, Heidelberg: Springer, pp. 264–280.
- Dey, Anind K.; Daniel Salber; Gregory D. Abowd; and Masayasu Futakawa (1999). The conference assistant: Combining context-awareness with wearable computing. *Digest of Papers. Third International Symposium on Wearable Computers, San Francisco, California, USA, 18–19 October 1999*. IEEE, pp. 21–28.
- Dong, Wei; Kate Ehrlich; Michael M. Macy; and Michael Muller (2016). Embracing cultural diversity: Online social ties in distributed workgroups. *CSCW'16. Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing, San Francisco, California, USA, 27 February – 2 March 2016*. New York: ACM, pp. 274–287.
- Farzan, Rosta; and Peter Brusilovsky (2007). Community-based conference navigator. *UM 2007. Proceedings of Workshop on Adaptation and Personalisation in Social Systems: Groups, Teams, Communities at the International Conference on User Modeling, Corfu, Greece, 25–29 July 2007*, Vol. 2007. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 30–39.
- Frydlinger, David; Jeanette Nyden; and Kate Vitasek (2013). Unpacking Collaboration Theory: What Every Negotiator Should Know to Establish Successful Strategic Relationships. <http://www.vestedway.com/vested-library>. Accessed 3 March 2019.
- Graves, Laura M.; and Priscilla M. Elsass (2005). Sex and sex dissimilarity effects in ongoing teams: Some surprising findings. *Human Relations*, vol. 58, no. 2, February 2005, pp. 191–221.
- Guy, Ido; and Luiz Pizzato (2016). People Recommendation Tutorial. *RecSys'16. Proceedings of the ACM Conference on Recommender Systems, Boston, Massachusetts, USA, 15–19 September 2016*. New York: ACM, pp. 431–432.
- Guy, Ido; Michal Jacovi; Adam Perer; Inbal Ronen; and Erel Uziel (2010). Same places, same things, same people?: mining user similarity on social media. *CSCW'10. Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work, Savannah, Georgia, USA, 06–10 February 2010*. New York: ACM, pp. 41–50.
- Heck, Tamara (2013). Combining social information for academic networking. *CSCW '13. Proceedings of the ACM Conference on Computer Supported Cooperative Work, San Antonio, TX, USA, 23–27 February 2013*. New York: ACM, pp. 1387–1398.
- Hoang, Dinh Tuyen; Van Cuong Tran; Van Du Nguyen; Ngoc Thanh Nguyen; and Dosam Hwang (2017). Improving Academic Event Recommendation Using Research Similarity and Interaction Strength Between Authors. *Cybernetics and Systems*, vol. 48, no. 3, March 2017, pp. 210–230.
- Hobman, Elizabeth V.; Prashant Bordia; and Cynthia Gallois (2004). Perceived dissimilarity and work group involvement: The moderating effects of group openness to diversity. *Group & Organization Management*, vol. 29, no. 5, October 2004, pp. 560–587.
- Hsiehchen, David; Magdalena Espinoza; and Antony Hsieh (2015). Multinational teams and dis-economies of scale in collaborative research. *Science advances*, vol. 1, no. 8: e1500211, September 2015, pp. 1–9.
- Jasny, Lorien; Joseph Waggle; and Dana R. Fisher (2015). An empirical examination of echo chambers in US climate policy networks. *Nature Climate Change*, vol. 5, no. 8, August 2015, pp. 782–786.
- Kahneman, Daniel (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, vol. 58, no. 9, September 2003, pp. 697.
- Kawakita, Ysuke; Shirou Wakayama; Hisakazu Hada; Osamu Nakamura; and Jun Murai (2004). Rendezvous enhancement for conference support system based on RFID. *SAINT 2004. International Symposium on Applications and the Internet Workshops, Tokyo, Japan, 26–30 January 2004*. IEEE, pp. 280–286.

- Knijnenburg, Bart P.; Martijn C. Willemsen; Zeno Gantner; Hakan Soncu; and Chris Newell (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, vol. 22, no. 4-5, October 2012, pp. 441–504.
- Kong, Xiangjie; Huizhen Jiang; Zhuo Yang; Zhenzhen Xu; Feng Xia; and Amr Tolba (2016). Exploiting publication contents and collaboration networks for collaborator recommendation. *PloS one*, vol. 11, no. 2: e0148492, February 2016, pp. 1–13.
- Lazarsfeld, Paul F; and Robert K. Merton (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, vol. 18, no. 1, pp. 18–66.
- Lee, Jae Kook; Jihyang Choi; Cheonsoo Kim; and Yonghwan Kim (2014). Social media, network heterogeneity, and opinion polarization. *Journal of communication*, vol. 64, no. 4, January 2014, pp. 702–722.
- Li, Baoli; and Liping Han (2013). Distance weighted cosine similarity measure for text classification. In: H. Yin, K. Tang, Y. Gao, F. Klawonn, M. Lee, T. Weise, B. Li, and X. Yao (eds.): *IDEAL 2013. International Conference on Intelligent Data Engineering and Automated Learning. Lecture Notes in Computer Science, Hefei, China, 20–23 October 2013*. Berlin, Heidelberg: Springer, pp. 611–618.
- Li, Lin; Anna Scaglione; Ananthram Swami; and Qing Zhao (2013). Consensus, polarization and clustering of opinions in social networks. *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, June 2013, pp. 1072–1083.
- Li, Jing; Feng Xia; Wei Wang; Zhen Chen; Nana Yaw Asabere; and Huizhen Jiang (2014). Acrec: a co-authorship based random walk model for academic collaboration recommendation. *WWW'14. Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, 07–11 April 2014*. New York: ACM, pp. 1209–1214.
- Mannix, Elizabeth; and Margaret A. Neale (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*, vol. 6, no. 2, October 2005, pp. 31–55.
- Marsden, Peter V. (1987). Core discussion networks of Americans. *American sociological review*, vol. 52, no. 1, February 1987, pp. 122–131.
- Mayer, Julia M.; Quentin Jones; and Starr Roxanne Hiltz (2015a). Identifying opportunities for valuable encounters: Toward context-aware social matching systems. *ACM Transactions on Information Systems (TOIS)*, vol. 34, no. 1, October 2015, pp. 1–32.
- Mayer, Julia M.; Starr Roxanne Hiltz; and Quentin Jones (2015b). Making Social Matching Context-Aware: Design Concepts and Open Challenges. *CHI'15. Proceedings of the ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18–23 April 2015*. New York: ACM, pp. 545–554.
- Mayer, Julia M.; Starr Roxanne Hiltz; Louise Barkhuus; Kaisa Väänänen; and Quentin Jones (2016). Supporting opportunities for context-aware social matching: An experience sampling study. *CHI'16. Proceedings of the ACM Conference on Human Factors in Computing Systems, San Jose, California, USA, 07–12 May 2016*. New York: ACM, pp. 2430–2441.
- McNee, Sean M.; John Riedl; and Joseph A. Konstan (2006). Being accurate is not enough. *CHI EA'06. Extended abstracts on Human factors in computing systems, Montréal, Québec, Canada, 22–27 April 2006*. New York: ACM Press, pp. 1097–1101.
- Mitchell, Rebecca; and Stephen Nicholas (2006). Knowledge creation in groups: The value of cognitive diversity, transactive memory and open-mindedness norms. *The Electronic Journal of Knowledge Management (EJKM)*, vol. 4, no. 1, December 2006, pp. 67–74.
- Mollica, Kelly A.; Barbara Gray; and Linda K. Treviño (2003). Racial homophily and its persistence in newcomers' social networks. *Organization Science*, vol. 14, no. 2, March – April 2003, pp. 123–136.
- Moody, James (2001). Race, school integration, and friendship segregation in America. *American journal of Sociology*, vol. 107, no. 3, November 2001, pp. 679–716.

- Moreland, Richard L.; and Robert B. Zajonc (1982). Exposure effects in person perception: Familiarity, similarity, and attraction. vol. 18, no. 5, September 1982, pp. 395–415.
- Muller, Michael; Kate Ehrlich; Tara Matthews; Adam Perer; Inbal Ronen; and Ido Guy (2012). Diversity among enterprise online communities: collaborating, teaming, and innovating through social media. *CHI'12. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA 5–10 May 2012*. New York: ACM, pp. 2815–2824.
- Nguyen, Tien T.; Pik-Mai Hui; F. Maxwell Harper; Loren Terveen; and Joseph A. Konstan (2014). Exploring the filter bubble: the effect of using recommender systems on content diversity. *WWW'14. Proceedings of the 23rd international conference on World wide web, Seoul, Korea, 7–11 April 2014*. New York: ACM, pp. 677–686.
- Nishibe, Yoshiyasu; Hiroaki Waki; Ichiro Morihara; Fumio Hattori; Toru Ishida; Toshikazu Nishimura; Hirofumi Yamaki; Takaaki Komura; Nobuyasu Itoh; Tadahiro Gotoh; Toyooki Nishida; Hideaki Takeda; Atsushi Sawada; Harumi Maeda; Masao Kajihara; and Hidekazu Adachi (1998). Mobile digital assistants for community support. *AI Magazine*, vol. 19, no. 2, June 1998, pp. 31.
- Olson, Gary M.; and Judith S. Olson (2000). Distance matters. *Human–computer Interaction*, vol. 15, no. 2-3, December 2000, pp. 139–178.
- Otsu, Nobuyuki (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, January 1979, pp. 62–66.
- Paasovaara, Susanna; Ekaterina Olshannikova; Pradthana Jarusriboonchai; Aris Malapaschas; and Thomas Olsson (2016). Next2You: a proximity-based social application aiming to encourage interaction between nearby people. *MUM'16. Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Rovaniemi, Finland, 12–15 December 2016*. New York: ACM, pp. 81–90.
- Parra, Denis; Wei Jeng; Peter Brusilovsky; Claudia López; and Shaghayegh Sahebi (2012). Conference Navigator 3: An online social conference support system. *UMAP 2012. UMAP workshops at the conference on User Modeling, Adaptation, and Personalization, Montréal, Québec, 16–20 July 2012*. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 1–4.
- Pu, Pearl; Li Chen; and Rong Hu (2011). A user-centric evaluation framework for recommender systems. *RecSys'11. Proceedings of the fifth ACM conference on Recommender systems, Chicago, Illinois, USA, 23–27 October 2011*. New York: ACM, pp. 157–164.
- Pu, Pearl; Li Chen; and Rong Hu (2012). Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction*, vol. 22, no. 4-5, October 2012, pp. 317–355.
- Rajagopal, Kamakshi; Jan M. van Bruggen; and Peter B. Sloep (2017). Recommending peers for learning: Matching on dissimilarity in interpretations to provoke breakdown. *British Journal of Educational Technology*, vol. 48, no. 2, February 2017, pp. 385–406.
- Rodan, Simon; and Charles Galunic (2004). More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness. *Strategic Management Journal*, vol. 25, no. 6, April 2004, pp. 541–562.
- Saldaña, Johnny (2015). *The coding manual for qualitative researchers*. London: SAGE.
- Sharma, Amit; and Dan Cosley (2016). Distinguishing between personal preferences and social influence in online activity feeds. *CSCW'16. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing, San Francisco, California, USA, 27 February – 02 March 2016*. New York: ACM, pp. 1091–1103.
- Sheldon, Michael R.; Michael J. Fillyaw; and W. Douglas Thompson (1996). The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiotherapy Research International*, vol. 1, no. 4, November 1996, pp. 221–228.
- Sie, Rory L. L.; Hendrik Drachsler; Marlies Bitter-Rijpkema; and Peter Sloep (2012). To whom and why should I connect? Co-author recommendation based on powerful and similar peers.

- International Journal of Technology Enhanced Learning (IJTEL)*, vol. 4, no. 1-2, August 2012, pp. 121–137.
- Sinha, Rashmi; and Kirsten Swearingen (2002). The role of transparency in recommender systems. *CHI EA'02. CHI'02 Extended Abstracts on Human Factors in Computing Systems, Minneapolis, Minnesota, USA, 20 - 25 April 2002*. New York: ACM, pp. 830–831.
- Stanovich, Keith E.; and Richard F. West (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, vol. 23, no. 5, October 2000, pp. 645–665.
- Terveen, Loren; and David W. McDonald (2005). Social matching: A framework and research agenda. *ACM transactions on computer-human interaction (TOCHI)*, vol. 12, no. 3, September 2005, pp. 401–434.
- Tsai, Chun-Hua; and Peter Brusilovsky (2016). A personalized people recommender system using global search approach. *ICConference 2016. Proceedings of IConference, Philadelphia, Pennsylvania, USA, 20–23 March 2016*. iSchools, pp. 1–5.
- Vassileva, Julita; Gordon McCalla; and Jim Greer (2003). Multi-agent multi-user modeling in I-Help. *User Modeling and User-Adapted Interaction*, vol. 13, no. 1-2, February 2003, pp. 179–210.
- Wongchokprasitti, Chirayu; Peter Brusilovsky; and Denis Parra-Santander (2010). Conference Navigator 2.0: community-based recommendation for academic conferences. *SRS '10. Workshop on Social Reminder Systems, Hong Kong, China 07–07 February 2010*. New York: ACM, pp. 1–5.
- Wuchty, Stefan; Benjamin F. Jones; and Brian Uzzi (2007). The increasing dominance of teams in production of knowledge. *Science*, vol. 316, no. 5827, May 2007, pp. 1036–1039.
- Ye, Teng; and Lionel P. Robert Jr. (2017). Does collectivism inhibit individual creativity?: The effects of collectivism and perceived diversity on individual creativity and satisfaction in virtual ideation teams. *CSCW'17. Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, OR, USA, 25 February – 01 March 2017*. New York: ACM, pp. 2344–2358.
- Yuan, Y. Connie; and Geri Gay (2006). Homophily of network ties and bonding and bridging social capital in computer-mediated distributed teams. *Journal of Computer-Mediated Communication*, vol. 11, no. 4, October 2006, pp. 1062–1084.
- Zaiane, Osmar R.; Jiyang Chen; and Randy Goebel (2007). DBconnect: mining research community on DBLP data. *WebKDD/SNA-KDD'07. Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, San Jose, California, 12 August 2007*. New York: ACM, pp. 74–81.
- Zenk, Lukas; Michael Smuc; and Florian Windhager (2014). Beyond the name tag. In: B. Lutz (ed.): *Wissen nimmt Gestalt an: Beiträge zu den Kremser Wissensmanagement-Tagen 2013*. Danube, Austria: Edition-Donau-Univ. Krems, pp. 215–224.
- Zhang, Amy X.; Anant Bhardwaj; and David Karger (2016). Confer: A Conference Recommendation and Meetup Tool. *CSCW'16 Companion. Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion, San Francisco, California, USA, 26 February – 2 March 2016*. New York: ACM, pp. 118–121.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.